REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	6/30/96	Progress - 6/1/95 - 5/31/96

4. TITLE AND SUBTITLE
Representation Issues in Signal Processing Applications

5. FUNDING NUMBERS
N00014-95-1-1163

6. AUTHOR(S)
Professor Avideh Zakhor

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The Regents of the University of California
c/o Sponsored Projects Office
336 Sproul Hall
Berkeley, CA 94720-5940

8. PERFORMING ORGANIZATION REPORT NUMBER
442427-23143

9. SPNSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
Ballston Tower One
800 North Quincy Street
Arlington, VA 22217-5660

10. SPONSORING /MONITORING AGENCY
Office of Naval Research
Seattle Regional Office

11. SUPPLEMENTARY NOTES
n/a

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release
Distribution Unlimited

12b. DISTRIBUTION CODE

19960805 032

13. ABSTRACT (Maximum 200 words)

We have developed an optimal bit allocation strategy for a joint source/channel video codec over noisy channel when the channel state is assumed to be known. Our approach is to partition source and channel coding bits in such a way that the expected distortion is minimized. The particular source coding algorithm we use is rate scalable and is based on 3D subband coding with multi-rate quantization. We have shown that using this strategy, transmission of video over very noisy channels still renders acceptable visual quality, and outperforms schemes that use equal error protection only. The flexibility of the algorithm also permits the bit allocation to be selected optimally when the channel state is in the form of a probability distribution instead of a deterministic state.

14. SUBJECT TERMS

15. NUMBER OF PAGES
8

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	

Summary of Activities for 6/1/1995 through 5/31/96

Avideh Zakhor
Department of Electrical Engineering and Computer Sciences
University of California
Berkeley, CA 94720

The advent of wireless personal communication services in recent years has created a number of challenging research problems in the areas of communications, signal processing and networking. A major challenge in dealing with the wireless channel has to do with its inherent unreliability. This is in contrast with wired networks in which the physical loss is very small, e.g. of the order of 10^{-9}.

It is generally argued upon that in practice, wireless video applications involve users moving at relatively slow speeds, rather than at tens of miles per hour. A direct consequence of slow moving users is that the resulting channels suffer from slow fading and shadowing effects. If the condition of this slowly changing channel is estimated, it is conceivable to adapt/modify the source coding, modulation, channel coding, power control, or any other aspects of transmission scheme to the channel condition. In particular, it is possible to vary the source and channel coding bit rates according to the channel condition, in such a way as to minimize the distortion of the received signals. Indeed, several researchers have applied this idea to speech [1, 7] and image [6, 2] transmission over the wireless links. In addition to adapting to channel conditions, one can protect different source bits using unequal error protection (UEP) schemes such as Rate Compatible Punctured Convolutional (RCPC) codes [5, 1, 2, 7, 6]. With the exception of [1] which deals with speech, the remaining papers mentioned above explicitly require the source coder to adapt to the channel condition. As an example, in [8] a whole new codebook might have to be designed and used in order to optimally match each new channel condition.

With the recent introduction of highly scalable video compression schemes such as [3], it is possible to generate one compressed bit stream, such that different subsets of it correspond to the compressed version of the same video sequence at different rates. Thus, if one uses such a source coder in the wireless scenario, there is no need to change the source coding algorithm,

or any of its parameters, as the channel conditions change. This is particularly attractive in multicast situations in heterogeneous networks where the wireless link is only a small part of a much larger network, and the source rate cannot be easily adapted to the individual receiver at the wireless node.

Over the past year, we have developed a technique for optimum partitioning of source and channel coding bits, for the scalable video compression algorithm described in [3]. By "optimum", we mean a partitioning which results in minimum expected value of distortion, which we choose to be Mean Squared Error (MSE). We will consider the case where the channel state information (CSI) is known, and the joint source/channel codec will adapt to the channel and optimally transmit video for the current channel state. In section 1, we briefly describe our source coding algorithm, section 2 formulates our problem and outlines our basic approach, section 3 contains details of our optimization based approach and section 4 includes results.

1 Scalable Source Coder

The source coder we use in this project is the scalable coder described in [3]. This coder has been shown to generate rates anywhere from tens of kilo bits to few mega bits per second with arbitrarily fine granularity. In addition, its compression efficiency has been shown to be comparable to standards such as MPEG-1 [3]. The fundamental idea behind it is to apply three dimensional subband coding to the video sequence to obtain a set of spatio temporal subbands. Subsequently, each subband coefficient is successively refined via layered quantization techniques. Finally, conditional arithmetic coding is applied to code different quantization layers. In doing so, the spatial, temporal, and inter-subband correlations, as well as correlation between quantization layers are taken into consideration to minimize the bit rate. The problem of optimal source allocation between different quantization layers of different subbands, in the absence of channel errors has been discussed in [3]. In the next section, we outline our approach for the case where channel errors are not zero.

2 Problem Formulation

The main problem we solve is: given a total number of bits C, and a given binary symmetric channel with bit error probability P_e, find the best source coding rate R_s and channel coding rate R_c such that $C = R_s + R_c$, and the expected value of MSE is minimized. This is equivalent to finding the optimal source to channel bit ratio R_s^o/R_c^o, with $R_s^o + R_c^o = C$, such that the distortion is minimized. To find these minima for various CSI's, our approach is to construct distortion curves $D(\frac{R_s}{C-R_s})$ and to locate the minima empirically.

In constructing the above curves for each specific value of $\frac{R_s^*}{C-R_s^*}$, we must

answer two questions: first, which quantization layers of which subband should be included in R_s? Second, to what extent should each one be protected? If n_k denotes the number of source bits in subband k of the source coder described in section 1, and $m_{i,k}$ denotes the number of channel bits used to protect source bit i of subband k, then our optimization problem is that of minimizing the expected value of MSE given by [9]:

$$
\begin{aligned}
E[MSE] &= \sum_{k=1}^{K} E[MSE_k] && (1) \\
E[MSE_k] &= \sum_{i=1}^{n_k} \left\{ \frac{g(m_{i,k})}{1-g(m_{i,k})} f_k(i) \prod_{j=1}^{i} [1-g(m_{j,k})] \right\} \\
&+ \prod_{j=1}^{n_k} [1-g(m_{j,k})] f_k(n_k+1) && (2)
\end{aligned}
$$

subject to the constraints:

$$\sum_{k=1}^{K} n_k \le R_s \qquad \sum_{k=1}^{K}\sum_{i=1}^{n_k} m_{i,k} \le R_c \tag{3}$$

In the above equations, MSE_k is the mean square error of subband k, K is the total number of subbands, $g(m_{i,k})$ is the bit error probability if $m_{i,k}$ channel bits are used for protection of bit i of subband k, and $f_k(i)$ is the resulting MSE if an error occurs at bit i of subband k, or equivalently the error sensitivity function.

To simplify the above problem, we can introduce a new variable m_k given by:

$$\sum_{i=1}^{n_k} m_{i,k} = m_k \tag{4}$$

to replace the variables $m_{i,k}$. The MSE of subband k can now be represented as a function of 2 variables n_k and m_k:

$$E[MSE] = \sum_{k=1}^{K} E[MSE_k(n_k, m_k)] \tag{5}$$

However, there is now an additional problem of optimally distributing m_k channel bits among n_k source bits, or the mapping from m_k to the set $\{m_{i,k}\}$. This is a discrete non-linear optimization problem which can be solved via the well known branch and bound technique. To speed up this optimization, we can impose the following pruning rules for the branch and bound algorithm: first, all bits within the same quantization layer must receive the same level of protection; second, higher quantization layers never receive more protection than lower quantization layers.

In the next section, we focus on the problem of optimizing 5 subject to the constraints in 3.

λ	μ	R_s	R_c
0.0287	0.095	64.0 kbits/s	67.3 kbits/s
0.0237	0.095	70.7 kbits/s	67.9 kbits/s
0.0187	0.095	72.4 kbits/s	67.9 kbits/s
0.0187	0.085	72.4 kbits/s	68.4 kbits/s
0.0187	0.075	72.4 kbits/s	69.3 kbits/s

Table 1: Source and Channel Rates for various λ and μ

3 OPTIMIZATION PROBLEM

Our solution to the optimization is based on a variation of Lagrange Multipliers, similar to the one developed in [4], with the exception that we are considering optimization of two sets of variables instead of one. The general theorem states the solution to the unconstrained optimization problem:

$$\min \left\{ \sum_{k=1}^{K} MSE_k(n_k, m_k) + \lambda \sum_{k=1}^{K} n_k + \mu \sum_{k=1}^{K} m_k \right\} \tag{6}$$

is also the solution to the constrained optimization problem described earlier, provided there exists λ and μ such that the constraints in 3 are met with equality. The questions left to solve this unconstrained problem are: (a) for given λ and μ, how to find the optimal solution to 6; (b) how to find λ and μ to meet the constraints stated in 3. We will answer those next.

3.1 Solving the unconstrained problem

We note that to solve 6 for given λ and μ, we can solve for each set of variables (n_k, m_k) independently. In other words, we can solve K independent minimization problems of the form:

$$\min \{MSE_k(n_k, m_k) + \lambda n_k + \mu m_k\} \tag{7}$$

To solve each optimization problem, we exploit the fact that the number of quantization layers for each subband takes on finitely many discrete values. For instance, the source video that is coded to 250 kilo bits per second, can have a maximum of 7 quantization layers in each subband. This implies that for each subband, n_k can take on one of 7 discrete values. Our approach to solving the optimization problem in 7 is exhaustive search. Specifically, we step through all values of n_k and for each one, find the optimum m_k using the branch and bound technique discussed in section 3.

3.2 Finding the Lagrange Multipliers

Note that for each set of (λ, μ), we can find the corresponding values of (n_k, m_k), and hence (R_s, R_c) using the approach outlined in section 3.1. As such, we can claim that R_s and R_c are both functions of both λ and

μ. To facilitate our goal of finding (λ, μ) that meet the constraints in 3 with equality, we make the simplifying assumption that R_s (R_c) is only a function of λ (μ). To show numerically that this is a valid assumption, we construct Table 1 for CSI of $Pe = 0.05$, using the approach described in section 3.1. We see that as we vary μ and keep λ constant, R_s stays relatively constant while R_c changes dramatically; the opposite is true when we keep μ constant and vary λ. This shows R_s is more sensitive to changes in λ than in μ, and the opposite is true for R_c, and our assumption is approximately valid. Another important observation one can make from Table 1 is that in general, the rate R_s (R_c) is inversely proportional to λ (μ). Putting these together, we can propose the following search strategy:

1. Guess an initial value for λ^1 and μ^1. Using these values, find corresponding R_s^1 and R_c^1 using procedures discussed in 4.1.
2. If target R_s^* is larger then R_s^1, let $\lambda^2 = \alpha\lambda^1$, where $0 < \alpha < 1$. Else $\lambda^2 = (1/\alpha)\lambda^1$ Similar procedure for μ^2.
3. Construct a linear function $R_s(1/\lambda)$ of variable $1/\lambda$ using previous two sets of points, $(\lambda^{t-2}, R_s^{t-2})$ and $(\lambda^{t-1}, R_s^{t-1})$. Target R_s^* by estimating λ^t using this linear function. Similar procedure for μ^t.
4. If R_s^t is within ϵ_s of R_s^* and R_c^t is within ϵ_c of R_c^*, stop; Otherwise, goto 3.

Empirically, we found $\alpha = 0.1$, $\epsilon_s = 0.05R_s^*$, and $\epsilon_c = 0.05R_c^*$ results in 15 to 20 iterations. Since we terminate the search algorithm when it reaches an approximate solution, a natural question is, "how far off is the answer from the ideal solution?". To answer this, we can apply the lemma in [4] and show that for our experimental results described in section 4, our algorithm terminates when it is within 5% MSE of the actual solution.

4 RESULTS

To test the above algorithm numerically, we combine the 3D scalable video codec and Rate-Compatible Punctured Convolutional Codes [5] to build our proposed joint source/channel codec. For source coding, we use 3 levels of spatial and 2 levels of temporal subband decomposition. We use 200 frames of the digitized video "raiders of the lost ark" to compute the distortion functions, and apply our bit allocation strategy to search for the optimal source to channel coding ratio R_s^o/R_c^o for various CSI ranging from 0.001 to 0.05. The total bit budget is $250kbits/s$. We see in Figure 1 that there exists a unique distortion minimum for various P_e.

We observe that as the error probability P_e increases, the total number of quantization layers selected decreases; this is due to the decrease of optimal source to channel ratio as the channel condition worsens. From $P_e = 0$ to $P_e = 0.001$, the layers that are dropped are mostly high frequency layers; this is due to the low error sensitivities of high frequency components. In

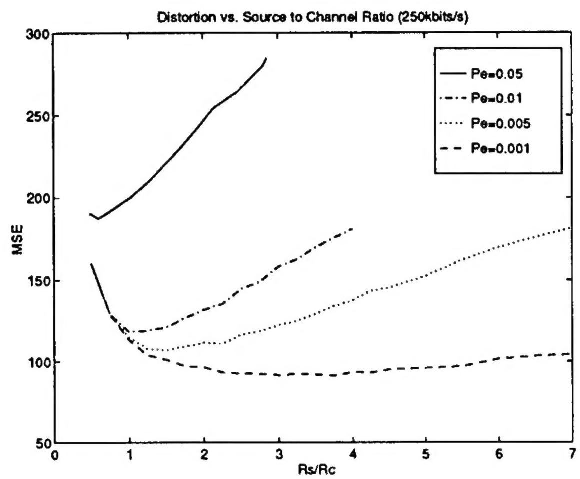

Figure 1: MSE vs. R_s/R_c for various CSI

poor channel condition, $P_e \geq 0.01$, the number of layers of low frequency subbands is reduced, resulting in a more uniform distribution of quantization layers among the subbands. This is because higher quantization layers are useless unless all the preceding lower layers are received error-free. Therefore a subband with too many quantization layers will render the higher layers frequently futile in poor channel condition.

Although we assume the knowledge of the channel, there are times when the estimate of the channel is incorrect. Using Figure 1, we can easily determine the performance degradations in such situations. For example, to find the approximate performance of the joint source/channel codec assuming $P_e = 0.01$ but operating at $P_e = 0.05$, we locate the point R_s^o/R_c^o for $P_e = 0.01$ on the curve $D_{0.05}(R_s/R_c)$.

If CSI is given in the form of a probability distribution function, then our proposed approach can be used to find the optimum operating point:

$$R_s^o/R_c^o = arg\left\{ \overset{\min}{R_s/R_c} \left[\sum_{P_{e'}\epsilon\Gamma} D_{P_{e'}}(R_s/R_c)P(CSI = P_{e'}) \right] \right\} \tag{8}$$

where Γ is the set of all possible CSI's. In this situation, our joint source/channel coding approach has better adaptation potential to the channel than previous schemes [2].

To show that our optimization strategy is essential in poor channel condition, $P_e = 0.05$, we compare its performance with other codecs in Figure 2. Curve a in Figure 2, shows the PSNR of the scalable codec under ideal noiseless conditions for 100 frames. The average PSNR in this case is 31.8 dB. Curve b in Figure 2 shows the PSNR of our proposed optimized codec operating at the optimal $R_s/R_c = 0.6$, with unequal error protection as described in earlier sections. The average PSNR in this case is about 4 dB lower than the ideal noiseless case. Curve c in Figure 2 shows the perfor-

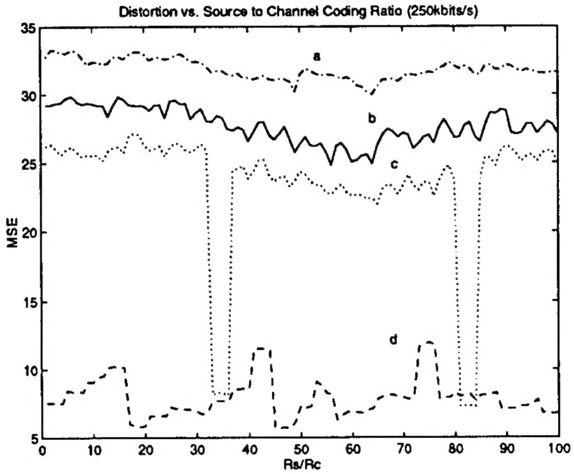

Figure 2: PSNR vs Frames for different Codecs: a) noiseless channel; b) proposed codec with unequal error protection $R_s^o/R_c^o = 0.6$, $P_e = .05$; c) equal error codec $R_s^o/R_c^o = 0.6$, $P_e = .05$; d) equal error codec $R_s/R_c = 2$, $P_e = .05$

mance of a codec operating at the optimal $R_s/R_c = 0.6$, but using equal error protection. This codec distributes R_s source bits using traditional bit allocation theory that assumes a noiseless channel, then channel codes these source bits with R_c channel bits equally. As seen, the PSNR is about the same as case b for most frames, except for occasional drops of 25 dB. These drops are a direct consequence of the fact that important source bits not being adequately protected and can result in objectionable degradations to the quality of the video. Finally curve d in Figure 2 shows the performance of the same equal error protection codec as in c but operating at non-optimal $R_s/R_c = 2$. As seen, the average PSNR of this codec is about 8 dB. The main conclusion to be drawn from Figure 2 is that optimal source/channel bit distribution does make a significant difference in both visual quality and PSNR.

References

[1] R.Cox, J.Hagenauer, N.Seshadri and C.Sundberg, "Subband Speech Coding and Matched Convolutional Channel Coding for Mobile Radio Channels," *IEEE Trans. SP*, vol.39, pp.1717-1731, August 1991.

[2] M.Ruf, J.Modestino, "Rate-Distortion Performance for Joint Source and Channel Coding of Images," *Proc. ICIP 1995*, pp.77-80, October 1995.

[3] D.Taubman and A.Zakhor, "Multirate 3-D Subband Coding of Video," *IEEE Trans. Image Proc.*, vol.3, pp.572-88, September 1994.

[4] Y.Shoham and A.Gersho, "Efficient Bit Allocation for an Arbitrary Set of Quantizers," *IEEE Trans. ASSP*, vol.36, pp.1445-1453, September 1988.

[5] J.Hagenauer, "Rate-Compatible Punctured Convolutional Codes (RCPC Codes) and their Applications," *IEEE Trans. Comm.*, vol.36, pp.389-399, April 1988.

[6] N. Tanabe, N. Farvardin, "Subband Image Coding Using Entropy-Coded Quantization over Noisy Channel," *IEEE Jour. Comm.* vol. 10, No. 5, pp.926-942, June 1992.

[7] H. Shi, P. Ho, and V. Cuperman, "Combined Speech and Channel Coding for Mobile Radio Communications", *IEEE Transactions on Vehicular Technology*, Vol. 43, No. 4, pp. 1078- 1087, November 1994.

[8] N. Farvardin, "A Study of Vector Quantization for Noisy Channels", *IEEE Transactions on Information Theory*, Vol. 36, Nol. 4, July 1990, pp. 799 - 809.

[9] G. Cheung, "Joint Source/Channel Coding of Scalable Video over Noisy Channels," Master's thesis, University of California at Berkeley, December 1995.